

How effective are social programs during conflicts? Instrumental variables evidence from the Angolan civil war*

Jean-Louis ARCAND[†]

The Graduate Institute, Geneva

Eric DJIMEU WOUABE

CERDI-CNRS, Université d'Auvergne

First draft: July 2008
This draft: January 2009

Abstract

How effective are social programs during conflicts? And how is the effectiveness of social programs affected by conflict intensity? In this paper, we consider the impact on child anthropometrics and household expenditures of the only major social program in Angola during the civil war. Our identification strategy is based on the political geography of program deployment. Our hypothesis is that one of the purposes of program implementation was to consolidate government political support in areas contested with UNITA, the main rebel group, as well to maintain the support of the population in areas to the rear of frontline communities. Based on a simple model of spatial competition, we show that the likelihood of treatment of a community should be increasing in the distance separating it from the government's forward base, and decreasing in the distance to UNITA's main base at the time. Using these exclusion restrictions to generate plausibly exogenous variation in treatment status, our linear instrumental variables estimates show that treatment at some point during the 1994-2000 period was associated with a 48.5% increase in household expenditures per adult equivalent, and a 0.359 increase in child height-for-age z -scores, in 2000. Results based on the local instrumental variables estimator show that each 1,000 additional deaths attributable to the civil war within a 20 km radius of the community shifts the marginal treatment effect associated with household income per adult equivalent up by 18.2%. This particular social fund was therefore associated with substantial benefits for treated communities, with these benefits being significantly increasing in the intensity of violence faced by the inhabitants.

Keywords: civil war, household expenditures, child anthropometrics, instrumental variables.

JEL Classification numbers: O19, H43, I12, I38.

*We thank Marie-Charlotte Buisson, Aude-Sophie Rodella, Henda Ducados and Victor Hugo Guilherme for very useful discussions. Jorge Dias Amaral provided an invaluable introduction to the Angolan literature. Lidia Galeano Germain was extremely helpful in allowing us to understand the IDR 2000 dataset. We thank the Bill and Melinda Gates Foundation, through a research grant administered by the Global Development Network (GDN), as well as the Angola 3rd Social Action Fund (FAS III) for financial support. The usual disclaimer applies.

[†]Corresponding author. Email: Jean-Louis.Arcand@graduateinstitute.ch.

1 Introduction

How effective are social programs during conflicts? And how is the effectiveness of social programs affected by conflict intensity? In this paper, we consider the impact on child anthropometrics and household expenditures of the only major social program in Angola during the civil war —the first two phases of the World Bank sponsored *Fundo de Apoio Social* (FAS I and II). Our identification strategy is based on the political geography of program deployment. Our hypothesis is that one of the purposes of program implementation was to consolidate government political support in areas contested with UNITA, the main rebel group, as well as to maintain the support of the population in areas to the rear of frontline communities. Based on a simple model of spatial competition in the tradition of Hotelling (1929), we show that the likelihood of FAS deployment should be an increasing function of the distance separating a given community from the government’s forward bases in Bailundo (Huambo province), and a decreasing function of the distance to UNITA’s main base at the time in Andulo (Bie province). Using these exclusion restrictions to generate plausibly exogenous variation in treatment status, and unique household data collected in Angola in 2000, our linear instrumental variables estimates show that treatment by the FAS during the 1994-2000 period was associated with a 48.5% ($se = 0.175$) increase in household expenditures per adult equivalent, and a 0.359 ($se = 0.166$) increase in child height-for-age z -scores. We then use the local instrumental variables estimator to explore the marginal treatment effect (MTE) of the FAS. In particular, we consider the interaction between treatment by the FAS and the intensity of conflict, using the remarkable data on episodes of violence during the Angolan civil war compiled by Ziemke (2007). In doing so, we find that each 1,000 additional deaths attributable to the civil war within a 20 km radius of the community shifts the MTE associated with household income per adult equivalent up by 18.2%. The FAS was therefore associated with substantial benefits for treated communities, with these benefits being significantly increasing in the intensity of violence faced by the inhabitants.

The remainder of this paper is organized as follows. Section 2 describes the functioning of the FAS, which is typical of World Bank-funded social funds operating in various parts of the world. It also summarizes the existing evidence on the impact of social funds, which is largely limited to their effects on educational outcomes, with relatively little attention having been devoted to household expenditures and child anthropometrics.

Section 3 presents our basic empirical specification, and demonstrates how, given targeting on communities that have unobservables that make them relatively poor, OLS-based estimates of the impact of the FAS on child anthropometrics and household expenditures are likely to be biased downwards, therefore calling for instrumental variables-based methods.

Section 4 presents our identification strategy and develops a theoretical model of the spatial division of political support between the government and UNITA, the main rebel group in Angola during the period of the civil war under consideration. We also present the household survey data, collected in 2000, on which our estimates will be based, and report the results of the first-stage reduced forms for treatment status. The latter do not reject the predictions of our theoretical model and appear to provide an appropriate degree of identification.

In section 5 we present our empirical results. We divide these into two parts. Section 5.1 presents linear instrumental variables (IV) results that show that treatment by the FAS was associated with highly significant and quantitatively large increases in household expenditures per adult equivalent and child height-for-age z -scores. The corresponding OLS-based estimates are negative and often statistically significant (depending upon whether covariates are included or not), highlighting the importance of adequately controlling for the endogeneity of treatment status. Section 5.2 tests for the presence of essential heterogeneity using the linearity test described (among other sources) in Heckman, Urzua, and Vytlacil (2006). We reject the absence of essential heterogeneity for household expenditures per adult equivalent, and do not for child height-for-age z -scores. Closer scrutiny of the household expenditures results using a carefully specified Roy model reveals a great deal of heterogeneity in the impact of treatment by the FAS, with an inverted U -shaped pattern emerging for the marginal treatment effect: the effect of treatment was negative (though marginally significant) for communities with unobservables that made them either highly likely or highly unlikely to receive treatment while, for communities with unobservables in between, the impact of treatment was positive and highly significant.

Section 6 combines our household survey data with extremely detailed conflict intensity data collated by Ziemke (2007). Continuing with the Roy model-based analysis of section 5.2, we show that, in communities which were subjected to particularly high levels of conflict, the MTE of the FAS was particularly large: an increase of 1,000 casualties within 20km of the community over the 1994-2000 period is associated with an increase of 18.2% ($se = 0.026$) in the marginal treatment effect of the FAS on household expenditures per adult equivalent. Section 7 concludes, and provides some tentative thoughts on the implications of our results for the deployment of social programs in situations of conflict.

2 The FAS and the impact of social funds

The FAS is an autonomous structure created by the Angolan government in October 1994. It is placed under the authority of the Ministry of Planning (MINPLAN) which determines its objectives. Between 1994 and 2000, the first two phases of the FAS (I and

II) were deployed in 9 out of the 18 Angolan provinces: Cabinda, Luanda, Kwanza Sul, Benguela, Huambo, Namibe, Huila, Cunene and Bengo. The areas of intervention of the FAS I and II are shown, along with the date of program deployment, on the map of Angola displayed in Figure 2. By and large, program deployment was confined to areas under government control, though certain provinces, such as Huambo, had only recently been retaken from UNITA, the main rebel group headed by Jonas Savimbi. An important purpose of the FAS was to aid in the transition towards peace (which, unfortunately, only lasted until 1996, when fighting broke out again until the final conclusion of the conflict in 2002), by targeting households that had been affected by the conflict and the ensuing economic crisis. The stated aims of the program were to improve community access to basic infrastructure, to improve the capacity of communities and local NGOs to plan, evaluate, manage and maintain community-level infrastructure, to create income generating activities both in rural and urban areas, and to contribute to an understanding of the sources of poverty in order to formulate more effective poverty reduction strategies.

The functioning of the FAS is similar to that of many social funds around the world. Local NGOs (*agencias de enquadramento*) work alongside communities (represented by the *nucleo comunitario*) so as to identify priority infrastructures and to prepare projects which are then submitted to the FAS. Provincial FAS offices vet submitted project proposals in coordination with the provincial government. All approved projects receive FAS funding, and their implementation is carried out by the communities themselves who must contribute 10% of total project costs, either in kind, through labor, or in cash. Between 1994 and 2001, the FAS funded \$29m worth of projects; 67% of these funds were provided by multilateral donors such as the World Bank, with the rest coming from bilateral donors and the Government of Angola (GOA).¹ A total of 685 projects were funded during the period, with 34% in education, 9% in health, 47.4% in water and waste management, and 9% in income generating activities. The average size of a FAS project is \$20,000.²

There is a growing literature on the impact of social funds though, with one exception, very little is known concerning their effects in an African context. Four papers, published in a special issue of the *World Bank Economic Review*, applied pipeline and propensity score matching methods (as well, to a more limited extent, difference-in-differences (DID), experimental and instrumental variables approaches) to evaluate the impact of social

¹We thank Victor Hugo Guilherme and Henda Ducados, respectively Director and Deputy-Director of the FAS III, for these numbers, which they presented at the World Bank Conference on Community-Driven Development, Ouagadougou, Burkina Faso, November 3-4, 2003.

²These numbers come from the FAS administrative database. The level of World Bank funding was far below that settled on in the IDA credit agreement with the Angolan government, which amounted to \$54m. The total level of FAS I and II funding is relatively small when compared with social funds in Latin America. Disbursements of the Bolivian SIF between 1994 and 1998 amounted to \$160m, those of the Nicaraguan FISE between 1991 and 1998 amounted to \$191m while disbursements of the Peruvian FONCODES between 1992 and 1998 totaled \$472m.

funds in Bolivia, Nicaragua, Armenia and Peru.

Evia, Newman, Pradhan, Ramiro, Rawling, and Ridder (2002) consider the impact of small-scale health, water and educational rural infrastructure funded by the Bolivian Social Investment Fund (SIF) over the 1993-1994 period. On the basis of experimental methods, they find that SIF-funded educational projects significantly improved the quality of school infrastructure (as measured by the number of books per student, class size, the presence of a latrine, clean drinking water and electricity in the school), while there was no impact on school enrollment rates, absenteeism, and test scores. The infrastructure did nevertheless significantly reduce dropout rates in primary classes by 3.8%. Propensity score matching- and DID-based evidence indicates that SIF-funded projects in health and water infrastructure increased usage and significantly reduced infant mortality rates.

Pradhan and Rawlings (2002) consider the impact of social infrastructure funded by the Nicaragua Emergency Social Investment Fund (FISE) between 1991 and 1998. They use pipeline and propensity score matching and show, based on household data collected in 1998, that educational projects significantly increased school enrollment, reduced absenteeism, and reduced the age of enrollment in the first grade. Health infrastructure investments were associated with greater usage of health centers, though there were no significant impacts on child anthropometrics. Similarly, FISE investments in water and waste management infrastructure had no discernible impact on child anthropometric outcomes.

Chase (2002) uses pipeline and propensity score matching methods to analyse educational and water infrastructure financed by the Armenian Social Investment Fund (ASIF) during the 1996-2000 period. Educational investments increased household expenditures on educational items, and increased enrollment rates. Water projects significantly increased access to potable water.

In contrast to the three preceding papers, where the results are almost always based on the assumption of selection on observables, Paxson and Schady (2002) consider the impact of the Peruvian Social Fund (FONCODES) over the 1993-1996 period using instrumental variables techniques. Using a community-level poverty index and the proportion of citizens having voted against the previous regime during the 1990-1993 period as exclusion restrictions, they show that FONCODES educational projects increased school enrollment rates for children between 6 and 11 years of age.

A fifth paper, also based on pipeline and propensity score matching methods, provides evidence on a social fund in Zambia (the MPU), a neighbor of Angola, which may be the most relevant in our case. Chase and Sherburne-Benz (2001) show that educational projects significantly increased school attendance, increased household educational expenses, whereas health projects, though they increased usage as well as vaccination rates, did not significantly affect health outcomes such as child diarrhea.

3 Empirical specification

Let i denote children, h households and c communities; N will denote sample size. The basic model that we are seeking to estimate is most easily understood to be given by:

$$Y_{ihc} = X_{ihc}\alpha + D_c\beta + \lambda_c\sigma + \varepsilon_{ihc}, \quad (1)$$

where Y_{ihc} is the $N \times 1$ vector associated with the outcome of interest (such as child health or household expenditures, in which case we shall write Y_{hc}), X_{ihc} is a $N \times K_X$ matrix of child, household and community control variables, D_c is an $N \times 1$ dummy that is equal to one when community c is treated by the FAS and zero otherwise, λ_c represents community level unobservables, and ε_{ihc} is a disturbance term. We assume that treatment status is given by a latent index model:

$$D_c^* = W_c\pi - \lambda_c = [X_c, Z_c]\pi - \lambda_c, \quad (2)$$

$$D_c = \mathbf{1}[D_c^* \geq 0], \quad (3)$$

where $W_c = [X_c, Z_c]$ is a $N \times K_W$ matrix of community-level covariates that determines treatment status, which we assume for the time being to be orthogonal with respect to λ_c and ε_{ihc} , $\mathbf{1}[\cdot]$ is the indicator function that is equal to one when the inequality in square brackets is satisfied and zero otherwise, X_c is a matrix of covariates that determines treatment status and that also appear in the outcome equation, and Z_c is a matrix of variables that determine treatment status but that do not appear in (1). The "minus" sign in front of λ_c in (2) is a standard notational convention. In this "common factor" specification, the scalar parameter σ in (1) determines the "degree of endogeneity" of treatment status. Our purpose is to consistently estimate β , which measures the impact of treatment by the social program on our outcome variable.

In the context of the implementation of social programs such as the FAS, program initiators will usually favor communities perceived to be particularly poor. If all community-level covariates that determine treatment status are included in X_{ihc} , then we are in a situation of "selection on observables" and the system we are seeking to estimate reduces to:

$$Y_{ihc} = X_{ihc}\alpha + D_c\beta + \varepsilon_{ihc}, \quad (4)$$

$$D_c^* = X_c\pi - \lambda_c, \quad (5)$$

$$D_c = \mathbf{1}[D_c^* \geq 0]. \quad (6)$$

Estimation of (4) by OLS will then provide an unbiased estimate of β .

On the other hand, if observables included in X_{ihc} do not adequately control for those factors that affect treatment status, community-level unobservables λ_c that also

determine treatment status will remain in (1), as parameterized by σ . If $\sigma > 0$, as is highly likely in the case of a poverty-reduction social program such as the FAS, the OLS estimate of β will be biased *downwards* because of the negative correlation between D_c and λ_c in (1). If this bias is sufficiently large relative to the true value of β , it may lead to the least squares estimate of β being *negative*. In this case, the only solution is estimation by instrumental variables. More precisely, we must identify a matrix Z_c of excluded instruments that (i) are a significant determinant of treatment status in (2), and (ii) can plausibly be excluded from (1).

4 Identification strategy

4.1 Intuition

The identification strategy that we adopt in this paper is based on the political geography underlying the deployment of the first two phases of the FAS. Our excluded instruments are given by the distance separating a given community (i) from the seat of the *União Nacional para a Independência Total de Angola* (UNITA) headquarters at the time of program initiation, and (ii) from the forward government bases which had just been captured in the context of an offensive in late 1994. More specifically, we will focus on the distance to the towns of Andulo (in Bie province) which served as UNITA military headquarters following the signing of the Lusaka Protocol in 1994 (see Spears 1999), which brought a temporary lull in the conflict and Bailundo (in Huambo province), a former UNITA base that constituted the forward prong of the government’s advance into the central highlands.³

Given that it is likely that one of the purposes of the government, through program deployment, was to maximize the political benefit from doing so, it is highly probable that the geographic location of a community with respect to Andulo and Bailundo was one of the factors determining the decision on where to deploy the FAS program. In particular, our hypothesis is that program deployment in areas located between government forward bases in Huambo province and UNITA headquarters in Andulo was geared towards gaining political support in these areas, while simultaneously maintaining support in the rear. We now turn to a spatial model that will help us to sharpen this intuition and provide a theoretical basis for our first-stage reduced forms.

³Other important UNITA headquarters were located at Lumbala N’guimbo in Moxico province, near the western border of Zambia, and at Jamba, in the southeastern corner of Cuando Cubango province, just north of the slice of Namibia known as the Caprivi strip. See Angola Peace Monitor (2000).

4.2 A spatial model of treatment status

Consider a spatial competition model in the tradition of Hotelling (1929).⁴ The government and the UNITA rebels are competing for political support in the area lying between the forward bases of the government in Bailundo (Huambo province), which was conquered in the 1994 offensive, and Andulo (Bie province), to which UNITA withdrew prior to the ceasefire brought about by the Lusaka Protocols. We assume that potential supporters are distributed uniformly along the line linking Bailundo to Andulo. The purpose of the government, through the geographical allocation of FAS program expenditures, is to maximize its level of support on the ground between a given community and Andulo, net of the cost of providing the program. Moreover, the government wishes to deploy the program in such a manner that it maintains support in areas "behind" the community; this includes areas located between Bailundo and the community, as well as areas located further back, and which are not directly contested by UNITA.

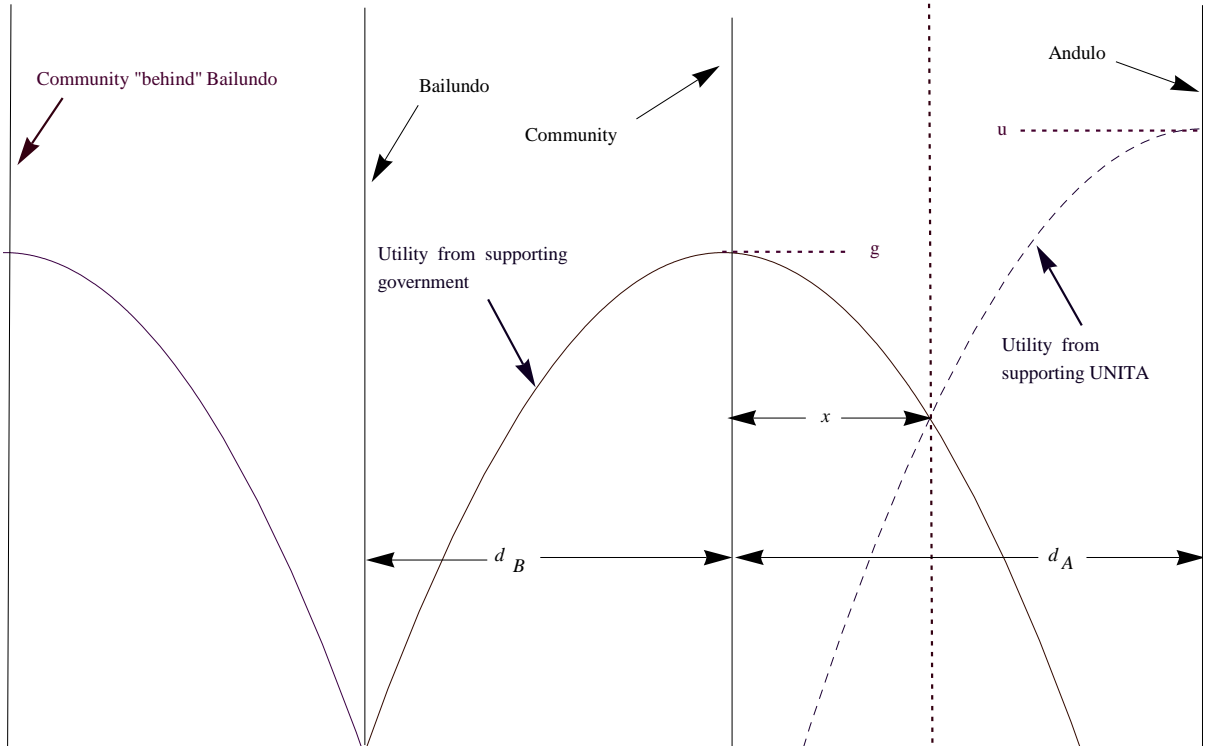


Figure 1: An illustration of the Hotelling spatial competition model underlying the first-stage reduced forms. If the government spends g in a given community, utility will be equal to $g - tx^2$ for a potential government supporter situated at distance x from that community; t parameterizes the degree of spatial differentiation.

Consider a community located somewhere between Andulo and Bailundo, at distance

⁴A more formal illustration of our identification strategy would be based on a model of spatial competition with two characteristics (in this case, longitude and latitude), such as Economides (1986).

d_A from Andulo, and d_B from Bailundo. Assume that the government spends g in the community in the context of the FAS program. For simplicity, we assume quadratic "transportation costs", in which an individual located at distance x from the community that receives the government expenditures has utility given by:

$$U = g - tx^2,$$

where t parameterizes the degree of horizontal differentiation.

A graphical illustration of the model is provided in Figure 1, where the inverted U-shaped utility functions reflect the quadratic transportation costs, which leads utility to be decreasing and concave as one moves away from the community being treated by the program. UNITA is located in Andulo. For areas lying between Andulo and the community treated by the FAS program, the dotted curve represents the utility of individuals who support UNITA. Given that the distance separating the community from Andulo is equal to d_A , an individual located at distance x from the community will be located at distance $d_A - x$ from Andulo.

An individual will be indifferent between supporting the government and UNITA when:

$$g - tx^2 = u - (d_A - x)^2,$$

where u is the level of utility of an individual located in Andulo proper and who supports UNITA. Solving the previous expression yields:

$$x = \frac{g - u + td_A^2}{2td_A}.$$

This is represented in Figure 1 by the intersection of the government's solid and UNITA's dotted utility curves. Individuals situated between the community and this intersection support the government, while those located between the intersection and Andulo support UNITA. When $g = u$, the population is evenly split between the government and UNITA. In Figure 1 we depict a situation in which $g < u$.

We assume that, for manifest security reasons, the government wishes to keep the political support of the population between the community in question and its bases in Bailundo. If we normalize the reservation level of utility of the consumer to zero (which corresponds to the horizontal axis in the Figure), such support between Bailundo and the community can be guaranteed by ensuring that:

$$g - td_B^2 \geq 0. \tag{7}$$

If the cost government expenditures is given by $\frac{1}{2}g^2$, where the convexity of the function reflects waste and the effect of rent-seeking activities, the government's optimization

problem is then given by:⁵

$$\max_{\{g\}} \frac{g - u + td_A^2}{2td_A} + d_B - \frac{1}{2}g^2 \quad s.t. \quad td_B^2 - g \leq 0,$$

which yields:

$$g^* = \begin{cases} \frac{1}{2td_A} & \text{for } \frac{1}{2t^2d_B^2} > d_A \\ td_B^2 & \text{otherwise} \end{cases}. \quad (8)$$

Note that for the government to maintain support amongst the population located between communities situated "behind" Bailundo and Bailundo itself, only the inequality given in (7) need be satisfied. In Figure 1 we depict a situation in which $\frac{1}{2td_A} = td_B^2$.

The implications of this highly stylized model of political competition are that FAS expenditures in a given community are: (i) *decreasing* in the distance d_A to Andulo in communities that are particularly close to this locality, (ii) *increasing* in the distance d_B to Bailundo for communities that are located far from Andulo. The model therefore suggests that the reduced form for treatment status, in which $Z_c = [d_A, d_B]$, should be given by:

$$D_c^* = X_c\pi_X + d_A\pi_A + d_B\pi_B - \lambda_c, \quad (9)$$

where $\pi_A < 0$ and $\pi_B > 0$. Whether our theoretical model provides the required identification, and whether it roughly corresponds to what is observed in the data can readily be tested by estimating (9) and examining both the significance and the signs of the coefficients associated with the distances separating each community from Andulo and Bailundo, respectively.

4.3 Data

The estimates presented in this paper are based on the 1999-2000 *Inquérito aos agregados familiares sobre despesas e receitas* ("national household survey on expenditures and incomes", henceforth, IDR 2000), which was collected in the provinces of Cabinda, Lunda, Lunda Norte, Benguela, Namibe, Huila and Cunene. Given the unstable security situation at the time, the survey is roughly representative of areas of Angola under effective government control. The IDR 2000 includes information on household composition, expenditures, and child health.⁶ It uses a stratified sampling design in which 12 households were surveyed in a random fashion in each *aldeia* (village) in rural areas and *bairro*

⁵Given that our goal is to characterize government spending behavior, we take u as being fixed. Allowing UNITA to choose u in an optimal fashion does not qualitatively modify the results of the model, since the expression for optimal government expenditures g^* given below in (8) is independent of UNITA behavior.

⁶We thank Lidia Galeano Germain for helping us understand the structure of the survey, as well as for filling in several gaps in the data. Formally speaking, the IDR data have not yet been officially released by the INE.

(neighborhood) in urban areas, in 50 *comunas*. Summary statistics for the two response variables and selected covariates for the 10,117 households and 8,328 children under the age of 5 included in the IDR 2000 data are presented in Table 1.⁷ The survey was carried out by the *Gabinete de monetarização das condições de vida da população, Instituto nacional de estatística* (INE), *Ministério do planeamento* (MINPLAN). An interesting analysis of the situation in Luanda using the IDR data is provided by Rodrigues (2003); the only other uses of the survey to date, to the best of our knowledge, are provided by UN (2002) and Tvedten (2002).⁸

The IDR data are particularly appropriate for testing the impact of the FAS, given that they correspond almost perfectly to the provinces in which there are communities treated by the FAS I and II. As such, our estimates will be essentially based on within-province differences between treatment and control communities, and will, if there are spillovers from treated to untreated communities, provide a lower bound on the impact of the program. Only Kwanza Sul and Bengo were treated by the FAS I and II and were not included in the IDR survey data, while only Lunda Norte is in the survey data without having any FAS-treated communities.

4.4 Treatment status

First-stage reduced forms of the determinants of treatment status corresponding to equation (9) are presented in the first two columns of Table 2. Column (1) presents linear probability estimates that correspond to the first-stage reduced form of the linear IV results for log household expenditures per adult equivalent that we present in section 5.1. Column (2) presents the corresponding first-stage reduced form for the child anthropometrics response variable. Column (3) presents the corresponding probit estimates that we will use in our analysis of the marginal treatment effect in section 5.2. We include the

⁷Other data collected in Angola during this period includes the *Inquérito Prioritário sobre as Condições de Vida aos Domicílios* (IPCVD), which was conducted in 1995, and focused only on urban areas, in Benguela, Cabinda, Catumbela, Luanda, Lobito, Lubango and Luena. It is therefore sometimes known as the "urban poverty study." For our purposes, however, it was carried out too soon after the initial deployment of the FAS program. It also has relatively limited geographical coverage. Another potential source of data is constituted by the *Inquérito sobre a Disposição e Capacidade no Pagamento dos Serviços Sociais Básicos* (IDCPSSB) which was carried out in 1998 in Huambo, Huíla, Luanda and Uíge. While the main focus was the ability and willingness of households to pay for basic social services, it also collected data on household budgets. According to UN (2002), the IDCPSSB was probably the most representative survey on household expenditure conducted to date, with 55% of the population sample in urban areas (including 24% in Luanda) and 45% in rural areas. The survey found an average per capita expenditure of \$2 a day. On these data, also see USAID (2002).

⁸A final source of microeconomic data for Angola during this period is provided by the 2001 *Multiple Indicator Cluster Survey*, (MICS 2001), carried out by UNICEF, which collected information on child anthropometrics, as well as on health and schooling issues, but no information on household budgets (a 1996 version of the survey also exists; see UNICEF (2003) for details on the most recent survey). Given the absence of household expenditure data, we have preferred to stick with the IDR. On child anthropometrics in Angola, apart from the MICS-based documents already cited, see GOA (2005). On educational attainment in Angola, see Bethke and Braunschweig (2003).

matrix of covariates X_{ihc} so as to avoid a "forbidden regression" problem in what follows. Covariates include child sex, age (in months) and age squared (for the child anthropometrics results), log household size, the dependency ratio, 8 dummies for the educational attainment of the household head, the household head's gender, 12 dummies for the type of water source used by the household (for the child anthropometric results), 12 dummies for the household head's language (which should control for ethnic-specific heterogeneity), six dummies for the household head's sector of employment, a rural-urban dummy, 7 provincial dummies, as well as the distance to Luanda and to the provincial capital. Standard errors are clustered at the *comuna* level in order to account for common shocks affecting all observations within a given *comuna*. As spelled out in our theoretical model in section 4.2, the distances of the community from Andulo and Bailundo are the two excluded instruments. Several aspects of the results are worth noting.

First, the likelihood of treatment by the FAS is a decreasing function of the distance of the community with respect to Andulo. Second, treatment by the FAS is an increasing function of the distance of the community with respect to Bailundo. Both of the point estimates are highly significant at the usual levels of confidence.⁹ Third, the results are in agreement with the predictions of our spatial model of program deployment suggesting, at the very least, that treatment status is not neutral with respect to the location of these two communities.

While it is of course possible that the location of a community with respect to Andulo and Bailundo could be correlated with unobservables that affect our response variables, this is highly unlikely in that neither of these locations is either particularly important in economic terms, on the one hand, or isolated, on the other. In both cases the direction of the ensuing bias in the IV estimates is uncertain, since it depends not only upon the covariances between treatment status and our exclusion restrictions (which are known from the results presented in Table 2), but also upon the sign of the covariance between our response variables and the distances separating a community from Andulo and Bailundo. Nevertheless, in order to minimize the eventuality of such location-induced biases, we include, amongst the included covariates, measures of the distance separating the community from Luanda (the national capital) and the corresponding provincial capital.

⁹Note that if we add the distance to Jamba (Cuando Cubango province) and the distance to Lumbala (Moxico province), two other locations of UNITA bases during this time period, as additional exclusion restrictions, the previously estimated coefficients remain statistically significant, with that associated with Bailundo doubling in size. The likelihood of treatment is increasing in the distance of the community from Jamba and decreasing in the distance from Lumbala. Note also that the linear IV results presented in section 5.1 remain both qualitatively and quantitatively unchanged when we use this extended instrument set. We prefer to confine ourselves to the more parsimonious instrument set since, in our opinion, our identification strategy is more transparent in this case. Results for the extended instrument set are, of course, available upon request.

5 Empirical results

5.1 Linear instrumental variables results

Table 3 presents instrumental variables estimates of the impact of treatment by the FAS on log household expenditures per adult equivalent and child height-for-age z -scores (HAZ). As shown by the coefficient reported in column (1), living in a FAS-treated community yields a 48.5% increase in household expenditures per adult equivalent, with an associated standard error of 0.175. The corresponding figure for total household income is 60.1% ($se = 0.232$). The coefficient reported in the uppermost portion of column (2) indicates that treatment by the FAS is associated with an increase of 0.359 in the HAZ ($se = 0.166$).

Both of these numbers are substantial, though the expenditures per adult equivalent result becomes understandable once the appallingly low initial level of income, caused in part by 25 years of civil war (and an additional 14 years of colonial-period struggles) is taken into account. The impact on child anthropometrics, which represents a reduction of 22.7% in the average height-for-age shortfall of the children included in the sample, is a relatively rare indication of the success of a social fund in terms of child health. On the other hand, the almost exclusive focus of existing research into social funds on educational outcomes (as summarized in section 2) may be in part to blame for this impression.

In the lower portion of the Table, we report the corresponding OLS-based coefficients. When no covariates are included, the coefficient reported in column (1) implies a *fall* of 47.4% in household expenditures per adult equivalent (which is significant at the 5% confidence level), while when covariates are included, the coefficient increases to -0.055 ($se = 0.085$). This configuration of IV and OLS point estimates is what one would expect if the model sketched in equations (1), (2) and (3) holds: the OLS coefficient is biased downwards because of the negative correlation between treatment status D_c and community level unobservables λ_c that also affect the response variable; adding covariates X_{ihc} that are correlated with treatment status reduces this bias by purging part of the correlation between D_c and λ_c .¹⁰ Finally, the putative "true" positive effect of treatment by the program is only uncovered once *exogenous* variation in treatment status is induced by appealing to excluded instruments. Much the same is true of the OLS estimates for the HAZ,

5.2 The marginal treatment effect of the FAS

An important question in the policy arena is how treatment by social programs such as the FAS is affected by the intensity of conflict. Does greater conflict intensity increase

¹⁰For the HAZ, both OLS coefficients are negative, but the results with covariates is statistically significant while without is not, at usual levels of confidence.

or reduce the benefits that accrue from treatment? The answer to this question is a key element in deciding whether social programs should be deployed in conflict situations, or whether they may actually be counterproductive, for example if they generate additional rents that the parties in conflict may fight over. Before examining this question, however, we must see whether the statistically significant treatment effects uncovered above using standard linear instrumental variables techniques are constant, or whether they vary with unobservables that determine treatment status. In order to do so, we appeal to the local instrumental variables (LIV) estimator of the marginal treatment effect pioneered by Heckman and Vytlacil (1999) and more fully described in Heckman and Vytlacil (2005) and Heckman, Urzua, and Vytlacil (2006)

The point of departure is a pair of potential outcome equations: Y_{1ihc} corresponds to the outcome when $D_c = 1$, whereas Y_{0ihc} corresponds to $D_c = 0$; all parameters and, most importantly, the disturbance terms, are allowed to differ by equation:

$$Y_{1ihc} = X_{ihc}\alpha_1 + \beta + \lambda_c\sigma_1 + \varepsilon_{1ihc}, \quad (10)$$

$$Y_{0ihc} = X_{ihc}\alpha_0 + \lambda_c\sigma_0 + \varepsilon_{0ihc}. \quad (11)$$

The system is completed by appending our latent index model given by (2) and (3). In particular, letting λ_c be distributed according to the standard normal density with *cdf* $\Phi(\cdot)$, the propensity score is then defined as:

$$P(W_c) = \Pr(W_c\pi \geq \lambda_c) = \Phi(W_c\pi),$$

and treatment status is determined by:

$$D_c = \mathbf{1}[\Phi(\lambda_c) \leq P(W_c)]. \quad (12)$$

Note that, by the *Probability Integral Transform*, we can, without loss of generality, arbitrarily normalize the disturbance term λ_c so that it is distributed according to the uniform density over the unit interval, implying that treatment occurs when $W_c\pi \geq U_{Dc}$, where $U_{Dc} \sim \text{Unif}[0, 1]$, thereby preserving our notation for the propensity score as $P(W_c)$.

In section 4.4 (see Table 2, columns (1) and (2)), we presented first-stage reduced forms of the determinants of treatment status using the linear probability specification that corresponds to the application of standard instrumental variables techniques. In column 3 of Table 2, we estimate (12), which corresponds to a probit specification. Results are very similar to those reported earlier in columns (1) and (2). In particular, the distances to Andulo and Bailundo remain highly significant determinants of treatment by the FAS, and the signs predicted by our theoretical model remain. Results are qualitatively unchanged if we replace the probit specification by a logit. Histograms for the estimated propensity scores, for treated and untreated observations, are presented in Figure 3: the

region of common support is given by the $[0.0296, 0.9917]$ interval.

In order to obtain an expression in terms of observables (Y_{1ihc} and Y_{0ihc} are never simultaneously observed for a given individual), note that the observable outcome Y_{ihc} can be expressed as:

$$Y_{ihc} = D_c Y_{1ihc} + (1 - D_c) Y_{0ihc}. \quad (13)$$

Substituting (10) and (11) into (13) then yields:

$$\begin{aligned} Y_{ihc} = & X_{ihc}\alpha_0 + D_c\beta + D_c X_{ihc}(\alpha_1 - \alpha_0) \\ & + \lambda_c\sigma_0 + \varepsilon_{0ihc} + D_c[\lambda_c(\sigma_1 - \sigma_0) + (\varepsilon_{1ihc} - \varepsilon_{0ihc})]. \end{aligned} \quad (14)$$

It is worth noting that this equation collapses to the usual case (given above by equation (1)) that can be addressed using standard instrumental variables techniques (as with the results presented in section 5.1), when $\alpha_1 = \alpha_0 = \alpha$, $\sigma_1 = \sigma_0 = \sigma$ and $\varepsilon_{1ihc} = \varepsilon_{0ihc} = \varepsilon_{ihc}$. This observation is the basis for Heckman's test for the presence of essential heterogeneity, to which we now turn.

Taking the conditional expectation of (14) yields:

$$E[Y_{ihc} | X_{ihc}, P(W_c)] = X_{ihc}\alpha_0 + P(W_c)X_{ihc}(\alpha_1 - \alpha_0) + K(P(W_c)), \quad (15)$$

where:

$$K(P(W_c)) = P(W_c)\beta + [\lambda_c(\sigma_1 - \sigma_0) + (\varepsilon_{1ihc} - \varepsilon_{0ihc}) | \Phi(\lambda_c) \leq P(W_c)] P(W_c).$$

In the absence of essential heterogeneity, $P(W_c)X_{ihc}(\alpha_1 - \alpha_0)$ drops out of equation (15) and $K(P(W_c))$ reduces to $P(W_c)\beta$, leaving a specification that is *linear* in the propensity score.

That this is definitely not the case for log household income per adult equivalent is illustrated in Panel A of Table 5, which presents a semiparametric estimate, over the entire unit interval, of $K(P(W_c))$ in equation (15), which allows for a completely general relationship between the outcome and the propensity score.¹¹ A nonparametric test of the linearity of $K(P(W_c))$ in $P(W_c)$ rejects the null with an extremely low p -value, as reported in Table 3. The same is true when we replace the semiparametric specification of $K(P(W_c))$ with a fourth order polynomial in the propensity score, and when we add the interaction terms ($P(W_c)X_{ihc}$) between the covariates and the propensity score.

In contrast, estimates for child height-for-age z -scores fail to reject the absence of essential heterogeneity, as illustrated in Panel B of Table 5, which yields essentially a straight line for $K(P(W_c))$. To the extent that our identification strategy is valid, this

¹¹We use a specification based on cubic thin plate splines with the smoothing parameter chosen automatically by restricted maximum likelihood (REML).

suggests that the results presented for child anthropometrics in column (2) of Table 3 provide a reasonably faithful picture of the impact of treatment status on height-for-age z -scores.

The null of the absence of essential heterogeneity being strongly rejected for household expenditures by our data implies that it is not clear what the linear instrumental variables results reported earlier actually represent. This is because, in the presence of essential heterogeneity, standard treatment parameters —such as the average treatment effect (ATE), the effect of treatment on the treated (TT) and the effect of treatment on the untreated (TUT)— differ. For child health, our results point towards all of these treatment parameters being equal. Estimating the marginal treatment effect (MTE) using LIV will allow us to distinguish between these different treatment parameters for household expenditures. It will also allow us to estimate, within the statistical structure given by this Roy model, how conflict intensity affects the impact of the FAS on household expenditures.

In order to do so, we begin by extracting the linear part of $E[Y_{ihc}|X_{ihc}, P(W_c)]$, using our estimate of (15), and purging it from Y_{ihc} . This yields:

$$\tilde{Y}_{ihc} = Y_{ihc} - X_{ihc}\hat{\alpha}_0 - P(W_c)X_{ihc}(\widehat{\alpha_1 - \alpha_0}).$$

We then obtain an estimate of $\frac{\partial \tilde{Y}_{ihc}}{\partial p_c}$ by regressing \tilde{Y}_{ihc} nonparametrically on $P(W_c)$ over the region of common support. The MTE is then given by:

$$\begin{aligned} \Delta^{MTE}(x_{ihc}, u_{Dc}) &= \left. \frac{\partial E[Y_{ihc}|X_{ihc} = x_{ihc}, P(W_c) = p_c]}{\partial p_c} \right|_{p_c = u_{Dc}}, \\ &= x_{ihc}(\widehat{\alpha_1 - \alpha_0}) + \left. \frac{\partial K(P(W_c) = p_c)}{\partial p_c} \right|_{p_c = u_{Dc}}, \\ &= x_{ihc}(\widehat{\alpha_1 - \alpha_0}) + \left. \frac{\partial \tilde{Y}_{ihc}(P(W_c) = p_c)}{\partial p_c} \right|_{p_c = u_{Dc}}. \end{aligned}$$

The nonlinear portion ($\frac{\partial K}{\partial p_c} = \frac{\partial \tilde{Y}_{ihc}}{\partial p_c}$) of the MTE for log household income per adult equivalent is represented by the solid curve in Figure 4 (the dotted curves correspond to the ± 2 standard error confidence bands).¹² The solid horizontal line in the figure corresponds to *minus* the linear portion of the MTE ($x_{ihc}(\widehat{\alpha_1 - \alpha_0})$), evaluated at the mean sample values of x_{ihc} . In terms of understanding the overall magnitude of the MTE, this horizontal line will therefore shift up or down depending upon the subset of observations over which $x_{ihc}(\widehat{\alpha_1 - \alpha_0})$ is evaluated. An intuitive manner of interpreting the figure is therefore that different subsets of observations correspond to different vertical intercepts of this horizontal line, with each intercept corresponding to a particular normalization.

¹²Note that derivative estimation is an ongoing area of research in nonparametric statistics. A good primer on the topic is provided by Newell and Einbeck (2007).

As is evident upon inspecting Figure 4, the MTE displays an inverted- U pattern. For both low and high values of u_{Dc} , the MTE is negative (though not statistically significant for low values of u_{Dc}), whereas it becomes positive for intermediate values of u_{Dc} . For households living in communities with unobservables such that it was highly likely that they received treatment by the FAS, which corresponds to low values of u_{Dc} (which in turn corresponds to low values of λ_c in the latent index model given by (2)), the MTE is negative and insignificant. For communities with unobservables such that it was highly *unlikely* that they received treatment by the FAS, the MTE is both negative and statistically significant at the usual levels of confidence. Figure 4 makes it clear that the substantial treatment effect estimated using linear IV techniques is attributable in large part to communities with u_{Dc} lying roughly in the $[0.15, 0.75]$ interval. Figure 4 also makes clear that while no formal targeting mechanism was set up in the context of program deployment, there was a very high probability of the FAS bypassing communities whose benefits would have been negative. On the other hand, unobservables making it highly likely that a community would be treated were associated with MTEs that were effectively zero.¹³

6 Conflict intensity and the marginal treatment effect

It is important to keep in mind that our purpose here is *not* to consistently estimate the impact of conflict intensity on our response variables.¹⁴ Instead, our purpose is to assess whether the impact of the FAS on our response variables is a function of conflict intensity. In the context of our estimation strategy, assessing the effect of conflict intensity on the MTE is extremely straightforward. To see why, reconsider equation (15), in which we explicitly include conflict intensity, denoted by C_c , as a covariate:

$$E[Y_{ihc} | X_{ihc}, P(W_c)] = X_{ihc}\alpha_0 + C_c\delta_0 + P(W_c)X_{ihc}(\alpha_1 - \alpha_0) + P(W_c)C_c(\delta_1 - \delta_0) + K(P(W_c)). \quad (16)$$

¹³Using the weights provided in Heckman (2001), we can derive the three standard treatment parameters by computing the appropriate weighted averages of our estimates of the MTE, though these results are not particularly informative, given the highly nonlinear aspect of the curve. Note that these are empirical versions of the treatment parameters, in that the region of common support is not the full unit interval. Given that there is a large portion of the empirical MTE that is negative when evaluated at the mean sample value of conflict intensity, the estimated ATE is negative (ATE = -0.067 , $se = 0.197$) though it is not statistically significant at the usual levels of confidence. The same is true for the TT (= -0.180 , $se = 0.212$) and TUT (= -0.147 , $se = 0.170$).

¹⁴This is done in a companion piece.

The marginal treatment effect is then given by:¹⁵

$$\Delta^{MTE}(x_{ihc}, u_{Dc}) = c_c(\widehat{\delta_1 - \delta_0}) + x_{ihc}(\widehat{\alpha_1 - \alpha_0}) + \left. \frac{\partial \tilde{Y}_{ihc}(P(W_c) = p_c)}{\partial p_c} \right|_{p_c = u_{Dc}}.$$

The impact of conflict intensity on the MTE is therefore given by $c_c(\widehat{\delta_1 - \delta_0})$, which we can directly measure through a semiparametric estimate of (16).¹⁶ An important point, first noted by Heckman and Vytlacil (1999), and that is in contrast to standard econometric approaches, is that X_{ihc} may be correlated with the disturbances in the potential outcome equations: in other words, X_{ihc} need not be “exogenous” in any conventional definition of that term.¹⁷

Our conflict intensity variable is based on the painstaking work of Ziemke (2007) who, based on archives, libraries and news agency files (a total of 186 sources from over twenty countries were involved), constructed a database of 9,216 individual battle and massacre events that took place in the Angolan war over a forty-one year period (1961-2002), as reported in the press.¹⁸ She provides the number of victims and the geographical coordinates for each event, allowing us, for each location included in the IDR 2000 dataset, to calculate the number of war-related casualties within a given radius of the community, in a given year. We considered four different radii: 1, 5, 10 and 20 kilometres, and present results that correspond to the 20km radius; results for the three other radii were qualitatively similar, with minor quantitative differences. Summary statistics for the conflict intensity variables are presented in the lower portion of Table 1.

In quantitative terms, the point estimate of $\widehat{\delta_1 - \delta_0}$ (0.182, $se = 0.026$), reported in the lowermost portion of Table 4, implies that each additional 1,000 casualties within 20 km of the community during the 1995-2000 period is associated with an 18.2% increase in the MTE of the FAS. These results indicate that while the impact of treatment by the FAS on household expenditures per adult equivalent was characterized by a great degree of heterogeneity driven by unobservables that determined treatment status, communities that bore the brunt of the fighting were the biggest beneficiaries.¹⁹ In the Angolan

¹⁵Of course, $\tilde{Y}_{ihc} = Y_{ihc} - X_{ihc}\hat{\alpha}_0 - C_c\hat{\delta}_0 - P(W_c)X_{ihc}(\widehat{\alpha_1 - \alpha_0}) - P(W_c)C_c(\widehat{\delta_1 - \delta_0})$ in this case.

¹⁶See Carneiro, Heckman, and Vytlacil (2003) who show, in exactly the same statistical setup, that individuals with good results on the Armed Forces Qualifying Test (AFQT) have their MTE from completing college (the response variable is log wage) shifted upwards.

¹⁷The key orthogonality assumption is that the three disturbance terms must be orthogonal with respect to the excluded instruments, *conditional* on the included covariates: $(\lambda_c\sigma_1 + \varepsilon_{1ihc}, \lambda_c\sigma_0 + \varepsilon_{0ihc}, \lambda_c) \perp Z_c \mid [X_{ihc}, X_{hc}, X_c]$.

¹⁸This work is part of the Armed Conflict Location and Events Dataset (ACLED).

¹⁹Note that we can find no specification in which the marginal impact of treatment status on HAZ is a statistically significant function of conflict intensity. For example, if we revert to standard IV techniques, allow for an interaction term between treatment status and conflict intensity, assume that both conflict intensity and the interaction term are jointly endogenous (and use the product of the distance variables as the required third exclusion restriction), the interaction term remains statistically indistinguishable from zero.

context, at least, greater conflict intensity was associated with an increase in the benefits that local populations derived from treatment by this particular social fund.

7 Concluding remarks

In retrospect, the inverted- U shaped MTE for household expenditures per adult equivalent is perhaps not so surprising. Assume for argument's sake that program initiators would have been able to identify communities with particularly large MTEs. In this case, and if the purpose of targeting had been to concentrate deployment on communities which had the most to gain from the program while avoiding communities with particularly low (or negative) MTEs, one would expect to see a *downward-sloping* MTE curve: communities with unobservables such that they would have been highly likely (unlikely) to be treated would have large (small) MTEs.

That this is *not* the case indicates two things. First, the program was not targeted so as to maximize economic gains. Second, given that maximizing economic gains was not the goal, something else was the driving force behind program deployment, and our hypothesis in this paper, crystallized in our identification strategy based on the political geography of the Angolan conflict, has been that it was political influence on the ground, particularly in areas that were on the frontline at the time of program initiation.

In policy terms, and while one must be cautious in generalizing from results based on a specific conflict situation, especially one as complex as Angola, our findings suggest at least three lessons. First, if our identification strategy is valid, and if, indeed, the significant positive effect of treatment by the program on child anthropometrics is neutral with respect to conflict intensity, there is no reason *not* to deploy social programs in areas of conflict when maintaining child health is the paramount goal.

Second, given that the marginal treatment effect of the program on household expenditures per adult equivalent is increasing in conflict intensity, there are good economic reasons for concentrating program deployment in areas where violence is (or has been) particularly severe, though this prescription must necessarily be tempered by the dangers to program staff that stem from such a form of geographical targeting.

Third, while it would be preferable on efficiency grounds to target the program so as to guarantee a downward-sloping MTE curve, thereby ensuring that it is highly likely that those communities with particularly large MTEs receive treatment, such a prescription is undoubtedly naive from the political standpoint: in situations of conflict, social programs necessarily become a tool associated with grassroots support in the field, and ignoring government political goals in the interests of efficiency may, ultimately, be counterproductive in terms of the welfare of the populations concerned.

References

- ANGOLA PEACE MONITOR (2000): “Jamba Falls to FAA Following Heavy Fighting in Border Region,” *Angola Peace Monitor*, VI(5), 18 January 2000.
- BETHKE, L., AND S. BRAUNSCHWEIG (2003): *Global Survey on Education in Emergencies. Angola Country Report*. Women’s Commission for Refugee Women and Children, New York, NY.
- CARNEIRO, P., J. J. HECKMAN, AND E. VYTLACIL (2003): “Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education,” processed, University of Chicago, The American Bar Foundation and Stanford University, July 19.
- CHASE, R. S. (2002): “Supporting Communities in Transition: The Impact of the Armenian Social Investment Fund,” *World Bank Economic Review*, 16(2), 219–240.
- CHASE, R. S., AND L. SHERBUME-BENZ (2001): “Household Effects of Community Education and Health Initiatives: Evaluating the Impact of the Zambia Social Fund,” The World Bank, processed, January.
- ECONOMIDES, N. (1986): “Nash Equilibrium in Duopoly with Products Defined by Two Characteristics,” *Rand Journal of Economics*, 17(3), 431–439.
- EVIA, J. . L., J. NEWMAN, M. PRADHAN, C. RAMIRO, L. B. RAWLING, AND G. RIDDER (2002): “An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund,” *World Bank Economic Review*, 16(2), 241–274.
- GOA (2005): “Estudo de Caso de Angola Sobre O Direito Humano À Alimentação Adequada,” Luanda, Angola.
- HECKMAN, J. J. (2001): “Micro Data, Heterogeneity and the Evaluation of Public Policy: Nobel Lecture,” *Journal of Political Economy*, 109(4), 673–748.
- HECKMAN, J. J., S. URZUA, AND E. VYTLACIL (2006): “Understanding Instrumental Variables in Models with Essential Heterogeneity,” *Review of Economics and Statistics*, 88(3), 389–432.
- HECKMAN, J. J., AND E. J. VYTLACIL (1999): “Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects,” *Proceedings of the National Academy of Sciences*, 96(8), 4730–4734.
- (2005): “Structural Equations, Treatment Effects and Econometric Policy Evaluation,” *Econometrica*, 73(3), 669–738.
- HOTELLING, H. (1929): “Stability in Competition,” *Economic Journal*, 39(153), 41–57.
- NEWELL, J., AND J. EINBECK (2007): “A Comparative Study of Nonparametric Derivative Estimators,” Proceedings of the 22th IWSM, Barcelona.
- PAXSON, C., AND N. R. SCHADY (2002): “The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru,” *World Bank Economic Review*, 16(2), 297–319.

- PRADHAN, AND L. B. RAWLINGS (2002): “The Impact and Targeting of Social Infrastructure Investments: Lessons from Nicaraguan Social Fund,” *World Bank Economic Review*, 16(2), 275–295.
- RODRIGUES, C. U. (2003): “Recomposição Social e Urbanização Em Luanda,” *Centro de Estudos Africanos Occasional Papers*, Lisbon, Portugal.
- SPEARS, I. A. (1999): “Angola’s Elusive Peace: The Collapse of the Lusaka Accord,” *International Journal*, 54(4), 562–581.
- TVEDTEN, I. (2002): “Angola 2001/2002 Key Development Issues and Aid in a Context of Peace,” *Report R 2002:8*, Chr. Michelsen Institute Development Studies and Human Rights, Oslo, Norway.
- UN (2002): *Angola: The Post-War Challenges. Common Country Assessment 2002*. United Nations System in Angola.
- UNICEF (2003): *MICS Multiple Indicator Cluster Survey. Assessing the Situation of Angolan Children and Women at the Beginning of the Millennium. Analytical Report*.
- USAID (2002): *Angola Transition and Development Assessment*. U.S. Agency for International Development, Washington, D.C.
- ZIEMKE, J. (2007): “From Battles to Massacres: An Analysis of Changing Conflict Patterns in Angola: 1961-2002,” University of Wisconsin-Madison.



Figure 2: The FAS I and II intervention areas, in purple. Orange provinces (Bie and Zaire) correspond to provinces included in the FAS III. Green provinces are currently being treated.

	mean	median	std.
Response variables: y_{ihc} or y_{hc}			
Child height-for-age (HAZ) z -score	-1.577	-1.620	1.433
Household expenditures per adult equivalent (Kwanzas)	390	249	477
Selected covariates: x_{ihc} , x_{hc} and x_c			
Child age (months)	28.09	28.00	17.10
Child is female	0.490	0.000	0.499
Household is Kimbundo	0.469	0.000	0.499
Years education of household head	4.537	5.000	2.056
Acces to water in the house	0.120	0.000	0.325
Household size	7.27	7.00	3.17
Distance to:			
Luanda, in 10 km	38	38	37
Provincial capital, in 10 km	8	3	13
Exclusion restrictions: z_c			
Distance to UNITA headquarters in Andulo, in 10 km	46	45	19
Distance to Bailundo, in 10 km	50	46	17
Conflict intensity: C_c			
Number of casualties, 1995-2000 (in thousands), within a radius of:			
1 km	0.016	0.000	0.044
5 km	0.020	0.000	0.048
10 km	0.038	0.000	0.059
20 km	0.102	0.025	0.140

Table 1: Descriptive statistics: response variables and covariates (8,328 observations for child-specific variables, 10,017 observations for household-specific variables).

Dependent variable:	Treatment status: D_c		
	Equals one if the community was treated by the FAS I or II programs during the 1994-2000 period, and zero otherwise		
	Linear probability model		Probit
	Log household expenditures per adult equivalent response variable	Child height-for-age z -score (HAZ) response variable	Log household expenditures per adult equivalent response variable
	(1)	(2)	(3)
Exclusion restrictions Z_c			
Distance of community from (in 10 km):			
Andulo (Bie)	-0.143 (0.057)	-0.151 (0.077)	-0.814 (0.257)
Bailundo (Huambo)	0.100 (0.029)	0.071 (0.034)	0.560 (0.190)
Covariates $[X_{hc}, X_c]$			
Distance from Luanda	0.066 (0.040)	0.073 (0.054)	0.033 (0.014)
Distance from provincial capital	-0.017 (0.014)	-0.056 (0.009)	-0.005 (0.013)
Child variables included	no	yes	no
Household variables included	yes	yes	yes
Provincial dummies included	yes	yes	yes
\overline{R}^2 or pseudo- \overline{R}^2	0.671	0.680	0.750
F - statistic	F_{9902}^{98} [p -value] = 209 [0.000]	F_{8252}^{75} [p -value] = 244 [0.000]	

Table 2: The determinants of treatment status; 10,000 observations in columns (1) and (3), 8,327 observations in column (2). Standard errors clustered at the community level in parentheses.

Dependent variable Y_{hc} or Y_{ihc}	Log household expenditures per adult equivalent (1)	Child height- for-age z -score (HAZ) (2)
Instrumental variables estimates		
Treated by the FAS: D_c	0.485 (0.175)	0.359 (0.166)
<hr/>		
H_0 : D_c exogenous (p -value)	0.001	0.000
H_0 : absence of essential heterogeneity (p -value)		
4th order polynomial	0.000	0.738
4th order polynomial + interaction terms	0.000	0.814
nonparametric	0.000	0.068
<hr/>		
Excluded IVs. Distance from:		
Andulo, Bailundo	yes	yes
<hr/>		
OLS estimates: covariates not included		
Treated by the FAS: D_c	-0.474 (0.237)	-0.128 (0.069)
<hr/>		
OLS estimates: covariate included		
Treated by the FAS: D_c	-0.055 (0.085)	-0.185 (0.082)
<hr/>		

Table 3: Instrumental variables estimates of the impact of treatment by the FAS I or II programs on log household expenditures per adult equivalent and child height-for-age z -scores; 10,000 observations in column (1), 8,327 observations in column (2); child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included as appropriate. Standard errors clustered at the community level in parentheses.

Dependent variable Y_{hc}	Log household expenditures per adult equivalent
Selected covariates $X_{hc} : \widehat{\alpha}_0$	
Urban household	0.834 (0.043)
Female head	-0.071 (0.018)
Household size	-0.517 (0.013)
Dependency ratio	-0.039 (0.008)
Distance from Luanda	0.002 (0.000)
Distance from provincial capital	0.002 (0.000)
Conflict intensity $C_c : \widehat{\delta}_0$	0.002 (0.012)
Covariates interacted with propensity score $P(W_c)X_{hc} : \widehat{\alpha_1 - \alpha_0}$	
Distance from provincial capital	0.002 (0.000)
Conflict intensity $P(W_c)C_c : \widehat{\delta_1 - \delta_0}$	0.182 (0.026)

Table 4: Semiparametric estimate of the conditional expectation of household expenditures per adult equivalent: linear portion; 10,000 observations. Standard errors in parentheses.

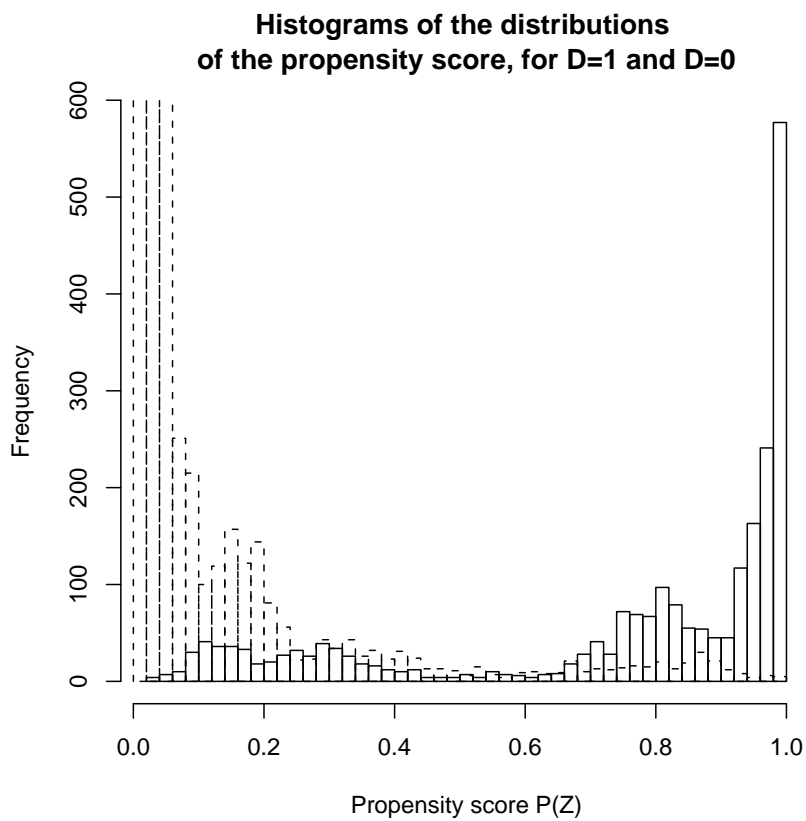
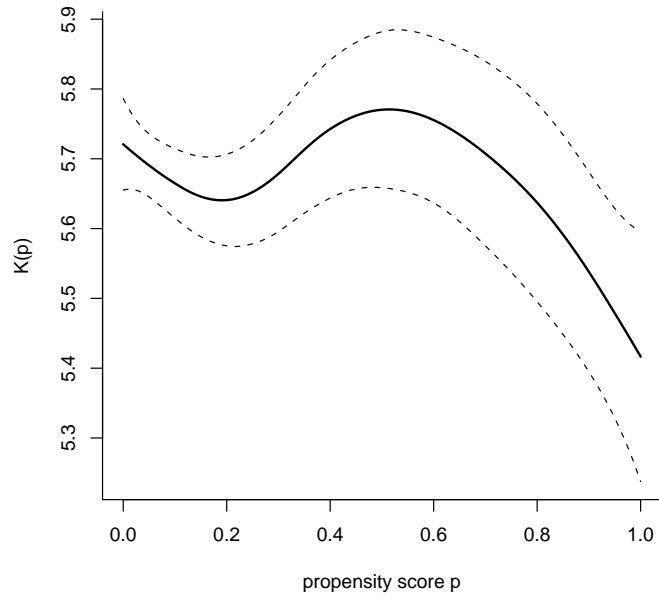
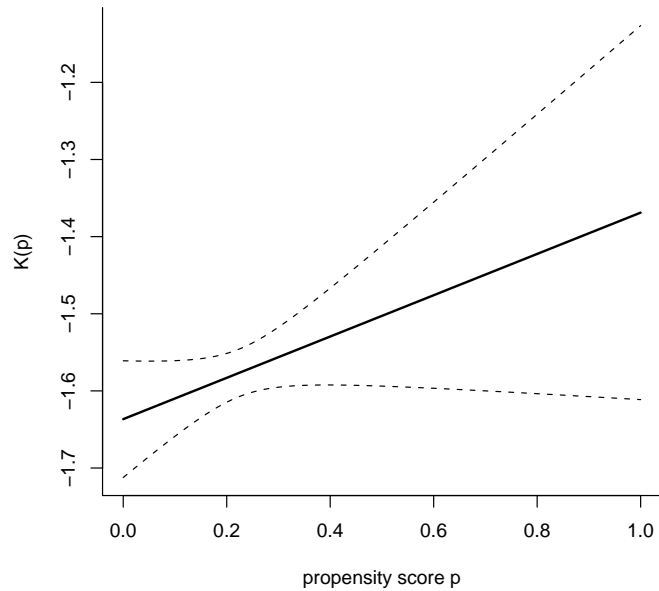


Figure 3: Histograms of estimated propensity score, for $D = 1$ (solid line) and $D = 0$ (dotted line). The region of common support is the $[0.0296, 0.9917]$ interval.



Panel A: Log household expenditures per adult equivalent.



Panel B: Child height-for-age z -score.

Table 5: The potential nonlinearity of $K(p)$ as a function of the propensity score, for log household expenditures per adult equivalent (Panel A, 10,000 observations), and child height-for-age z -scores (Panel B, 8,327 observations) (dotted lines represent plus/minus 2 standard error confidence bands).

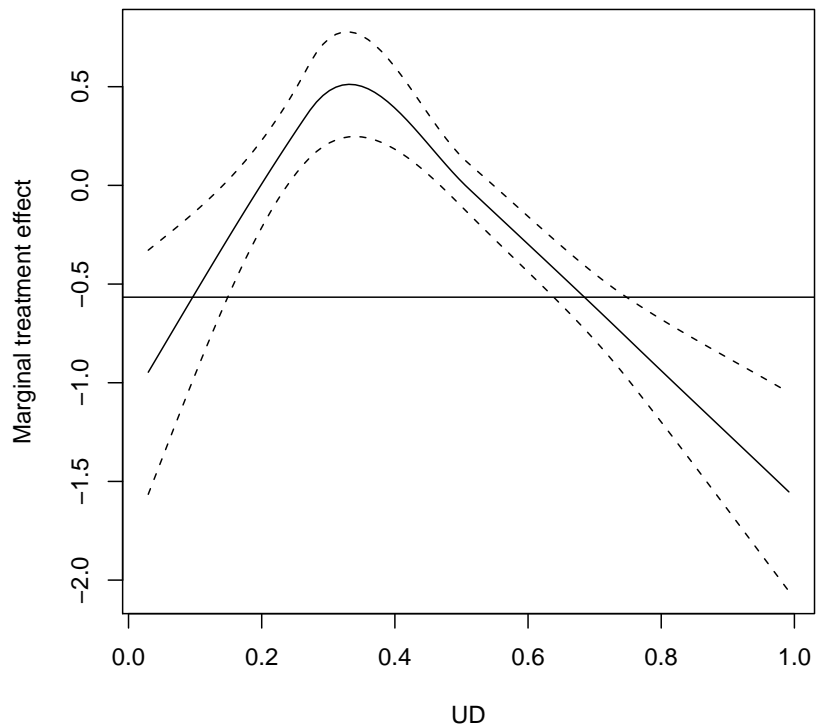


Figure 4: The marginal treatment effect of the FAS I and II on log household expenditures per adult equivalent (with ± 2 standard error confidence bands). Solid horizontal line corresponds to $-\widehat{c}_c(\widehat{\delta}_1 - \widehat{\delta}_0) - \widehat{x}_{ihc}(\widehat{\alpha}_1 - \widehat{\alpha}_0)$, the linear portion of the MTE evaluated at its sample mean.